Subgame perfection in Positive Recursive Games with perfect information

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 35, No. 1, February 2010, pp. 193–207
ISSN 0364-765X | EISSN 1526-5471 | 10 | 3501 | 0193
DOI 10.1287/moor.1090.0437

Subgame Perfection in Positive Recursive Games with Perfect Information

J. Flesch
Department of Quantitative Economics, Maastricht University, 6200 MD, Maastricht, The Netherlands,
j.flesch@maastrichtuniversity.nl

J. Kuipers, G. Schoenmakers, K. Vrieze
Department of Knowledge Engineering, Maastricht University, 6200 MD, Maastricht, The Netherlands
{kuipers@maastrichtuniversity.nl, gm.schoenmakers@maastrichtuniversity.nl, koos@ictvrieze.nl}

We consider a class of n-player stochastic games with the following properties: (1) in every state, the transitions are controlled by one player; (2) the payoffs are equal to zero in every nonabsorbing state; (3) the payoffs are nonnegative in every absorbing state. We propose a new iterative method to analyze these games. With respect to the expected average reward, we prove the existence of a subgame-perfect ε-equilibrium in pure strategies for every $\varepsilon > 0$. Moreover, if all transitions are deterministic, we obtain a subgame-perfect 0-equilibrium in pure strategies.

Key words: stochastic games; perfect information games; recursive games; subgame-perfect equilibria
MSC2000 subject classification: Primary: 91A15; secondary: 91A06
OR/MS subject classification: Primary: games/group decisions; secondary: stochastic
History: Received November 21, 2008; revised July 17, 2009, and November 2009. Published online in *Articles in Advance* January 27, 2010.

1. Introduction. In the theory of average reward stochastic games (with finite state and action spaces), a major open problem is whether or not ε-equilibria exist for all $\varepsilon > 0$. The famous game called the Big Match, which was introduced by Gillette [6] and solved by Blackwell and Ferguson [2], and the game in Sorin [21] demonstrated that 0-equilibria do not necessarily exist, and moreover, that history-dependent strategies are indispensable for establishing ε-equilibria. For two-player zero-sum games, Mertens and Neyman [12] showed the existence of ε-equilibria, in terms of ε-optimal strategies. Later, Vieille [24,25] provided a proof that ε-equilibria exist in all two-player stochastic games. For more than two players, however, only partial results are available, under rather restrictive conditions. For results on the existence of ε-equilibria in special classes, we refer to Thuijsman and Raghavan [22], Solan [17], Solan and Vieille [19], Simon [16], and Flesch et al. [3–5].

About subgame-perfect ε-equilibria even less is known. Lately, the class $\mathcal{G}$ of n-player recursive games with perfect information has received considerable attention. Here, recursive refers to the property that the payoffs are equal to zero in all nonabsorbing states, whereas perfect information means that, in any state of the game, at most one player has more than one action (this player controls the state). Let $\mathcal{G}^+$ denote the subclass of games in $\mathcal{G}$ in which all the payoffs are nonnegative in all absorbing states. In $\mathcal{G}^+$, the players have an incentive to eventually reach an absorbing state. It is known that subgame-perfect 0-equilibria do not always exist in $\mathcal{G}$ (cf. Solan and Vieille [20]) and also that stationary strategies are, in general, not sufficient for subgame-perfect ε-equilibria in $\mathcal{G}^+$ (cf., for example, Kuipers et al. [7], or Example 4 below).

We remark that, although Thuijsman and Raghavan [22] showed the existence of 0-equilibria for all games in class $\mathcal{G}$, their construction is not subgame perfect, in general, because it involves punishments, where $n-1$ players minimize a deviating player without regard to their own payoffs.

Solan [18] showed for a subclass of games in $\mathcal{G}$ that subgame-perfect ε-equilibria exist for all $\varepsilon > 0$. His proof requires two restrictions for games in $\mathcal{G}$: (1) each player controls exactly one nonabsorbing state and (2) each player in his own state has the choice between two actions: one *absorbing* action leading immediately to an absorbing state, and a *nonabsorbing* one leading to the n nonabsorbing states according to a probability distribution. It is essential that this probability distribution is the same for all n nonabsorbing actions. His proof is based on the analysis of a specific type of differential inclusions. A generalization of this result to periodic probability distributions can be found in Mashiah-Yaakovi [11].

For a subclass of games in $\mathcal{G}^+$, called free transition games, Kuipers et al. [7] proved the existence of subgame-perfect 0-equilibria in pure strategies. A game $G \in \mathcal{G}^+$, with T denoting the set of nonabsorbing states of G, is called a free transition game if it satisfies the following condition: in any state $s \in T$, the action space of the controlling player i_s is exactly $\{0\} \cup T$, where action 0 is *absorbing* and leading immediately to an absorbing state, and where action $a \in T$ is *nonabsorbing* and leads to nonabsorbing state a with probability 1.

In this paper, we prove the existence of subgame-perfect ε-equilibria in pure strategies, for all $\varepsilon > 0$, within the whole class $\mathcal{G}^+$. For the subclass of games with deterministic transitions, the proof simplifies, and we obtain subgame-perfect 0-equilibria in pure strategies. We introduce an inductive method to analyse these games. Although this method relies on the properties of the class $\mathcal{G}^+$, it is very natural and may initiate further results outside this class.[1]

We also wish to mention two related models, in which there is an ongoing investigation regarding the existence of equilibria and subgame-perfect equilibria. First, we mention a class of Dynkin games, where in each state (of a possibly infinite state space), the controlling player has one nonabsorbing action, called "continue," and one absorbing action, called "quit" (cf., for example, Solan and Vieille [20], Solan [18], or Mashiah-Yaakovi [10,11]). Another related model is the class of so-called stopping games, where in contrast with Dynkin games, several players may choose simultaneously between actions "continue" and "quit" (cf., for example, Shmaya et al. [15], Shmaya and Solan [14]).

For a coalitional game model where the transitions are deterministic, Vartiainen [23] independently proved a result that is similar to ours. He formulates an equilibrium concept requiring that an active coalition or any subcoalition thereof cannot benefit by deviating from a proposed well-defined strategy that specifies, for each history of coalitional moves, an active coalition and its move. Here, "well-defined" means that the strategy eventually leads to an outcome for which no further move is prescribed. Thus, this equilibrium concept applied to a coalitional game model resembles the notion of subgame-perfect equilibria applied to positive recursive games with deterministic transitions. Vartiainen [23] proves that a strategy satisfying his equilibrium concept exists. Moreover, his proof is based on the iterative application of majorization operations, similar to the iterative scheme presented in this paper.

This paper is structured as follows. In §2, we present the model and the main results together with the main idea behind the proof. In §3, we prove a useful lemma claiming that, for the main theorem, it suffices to consider games in a subclass $\tilde{\mathcal{G}}^+$ of $\mathcal{G}^+$. In §4, we introduce notions that will play an important role in our analysis. Finally, in §5, we present the proof of the main theorem for the subclass $\tilde{\mathcal{G}}^+$.

2. The model and the main results.

The class $\mathcal{G}^+$ of stochastic games. An n-player stochastic game in class $\mathcal{G}^+$ is given by (1) a nonempty set of players $N = \{1, \dots, n\}$, (2) a nonempty and finite set of states S, (3) for each state $t \in S$, an associated *controlling* player $i_t \in N$, (4) for each state $t \in S$, a nonempty and finite set of actions A_t, (5) for each state $t \in S$ and each action $a \in A_t$, a transition probability distribution $p_t(a) = (p_t(a, u))_{u \in S}$, (6) for each state $t \in S$ and each action $a \in A_t$, a payoff $r_t^i(a) \in \mathbb{R}$ to each player i such that the payoffs are equal to 0 in all non-absorbing states and the payoffs are nonnegative in all absorbing states. Here, a state t is called absorbing if $p_t(a, t) = 1$ for all actions $a \in A_t$; otherwise t is called nonabsorbing.

The game is to be played at stages in $\mathbb{N}$ in the following way. At any stage m, in the present state $s_m \in S$, the controlling player i_{s_m} has to choose an action a_m from the action set A_{s_m}. The chosen action a_m induces a payoff $r_{s_m}^j(a_m)$ to each player j, and a transition to a new state according to the transition probability distribution $p_{s_m}(a_m)$, where play will continue at stage $m+1$. We assume complete information (i.e., the players know all the data of the stochastic game), full monitoring (i.e., the players observe the present state and the action chosen by the controlling player), and perfect recall (i.e., the players remember all previous states and actions). The game starts in an initial state $s \in S$.

Strategies. A mixed action in state $t \in S$ for player i_t is a probability distribution on A_t. The set of these mixed actions is denoted by X_t. For $a \in A_t$, let $S_t(a) = \{u \in S \mid p_t(a, u) > 0\}$ denote the set of states to which transition occurs with a positive probability when action a is taken at state t. Let $H_{s,t}$ denote the set of all possible sequences $(s = s_1, a_1, \dots, s_m, a_m, s_{m+1} = t)$ of arbitrary but finite length, where for every $k = 1, \dots, m$ we have that (1) s_k is a state and a_k is an action of the controlling player i_{s_k} in state s_k, (2) $s_{k+1} \in S_{s_k}(a_k)$. Thus, $H_{s,t}$ is the set of all possible histories starting in initial state s and ending in state t.

A strategy π^i for player i and initial state s is a decision rule that, for any history $h \in H_{s,t}$ with $i = i_t$, prescribes a mixed action $\pi^i(h) \in X_t$. We use the notation Π_s^i for the set of strategies for player i and initial state s.[2] A strategy $\pi^i \in \Pi_s^i$ is called pure if every prescription $\pi^i(h)$ places probability 1 on one action. If the

[1] One of the referees suggested an approach, which is based on the main result, to prove the existence of subgame-perfect ε-equilibria in all perfect information stochastic games, with finite state and action spaces, when each player evaluates his sequence of payoffs by taking the limit superior.

[2] In this paper, a strategy provides prescriptions for histories with a given initial state, and it provides no prescriptions for other initial states. This is to conform with the concept of a continuation strategy, to be defined later on.

mixed actions prescribed by a strategy only depend on the final state of h, then the strategy is called stationary. We use the notation Π_s for the set of joint strategies $\pi = (\pi^i)_{i\in N}$ with $\pi^i \in \Pi_s^i$ for $i \in N$. A joint strategy $\pi = (\pi^i)_{i\in N}$ is pure if π^i is pure for all $i \in N$, and it is stationary if π^i is stationary for all $i \in N$.

Consider a strategy $\pi^i \in \Pi_s^i$ and a history $h \in H_{s,t}$. The continuation strategy $\pi^i[h] \in \Pi_t^i$ for player i and initial state t prescribes mixed actions for histories $h' \in H_{t,u}$ with $i_u = i$ according to π^i, but as if h had happened before h'. More formally, the continuation strategy $\pi^i[h]$ prescribes the mixed action $\pi^i[h](h') = \pi^i(h \oplus h')$ in state u. Here, $h \oplus h'$ is the history obtained by concatenation of h and h' (where t, the final state of h and the initial state of h', merge to one state t in $h \oplus h'$). We use the notation $\pi[h]$ to denote the joint continuation strategy, associated with $\pi = (\pi^i)_{i\in N}$ and $h \in H_{s,t}$.

Rewards. For initial state $s \in S$ and a joint strategy $\pi \in \Pi_s$, the sequences of payoffs are evaluated by the expected average reward, which is given for player i by

$$\phi^i(\pi) := \liminf_{M\to\infty} \mathbb{E}_\pi\left(\frac{1}{M}\sum_{m=1}^{M} R_m^i\right) = \liminf_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} \mathbb{E}_\pi(R_m^i),$$

where R_m^i is the random variable for the payoff for player i at stage m, and where $\mathbb{E}_\pi$ stands for expectation with respect to play according to the joint strategy π.

Equilibria. A joint strategy $\pi = (\pi^i)_{i\in N} \in \Pi_s$ is called a (Nash) ε-equilibrium for initial state s for some $\varepsilon \geq 0$ if

$$\phi^i(\sigma^i, (\pi^j)_{j\in N-\{i\}}) \leq \phi^i(\pi) + \varepsilon \quad \forall\, \sigma^i \in \Pi_s^i,\ \forall\, i \in N,$$

which means that no player can gain more than ε by a unilateral deviation. A strategy profile $\pi = (\pi_s)_{s\in S}$, with $\pi_s \in \Pi_s$ for all $s \in S$, is called a (Nash) ε-equilibrium for some $\varepsilon \geq 0$ if π_s is an ε-equilibrium for initial state s for all $s \in S$. As we mentioned in the introduction, a 0-equilibrium exists in every game in the class $\mathscr{G}^+$ (even in $\mathscr{G}$), but not in stationary strategies in general.

Subgame-perfect equilibria. A joint strategy $\pi \in \Pi_s$ is called a subgame-perfect ε-equilibrium for initial state s for some $\varepsilon \geq 0$ if for any $t \in S$ and history $h \in H_{s,t}$, the joint continuation strategy $\pi[h]$ is an ε-equilibrium for initial state t. A strategy profile $\pi = (\pi_s)_{s\in S}$ with $\pi_s \in \Pi_s$ for all $s \in S$, is called a subgame-perfect ε-equilibrium for some $\varepsilon \geq 0$ if π_s is a subgame-perfect ε-equilibrium for initial state s for all $s \in S$. Subgame perfection is a refinement of the equilibrium concept.

Our main results concern the existence of subgame-perfect ε-equilibria in the class $\mathscr{G}^+$ of stochastic games. We will pay special attention to the case when all transitions are deterministic, i.e., when for every state $s \in S$ and action $a \in A_s$, there exists a state $t \in S$ such that $p_s(a, t) = 1$.

Main Theorem. *Every stochastic game G in class $\mathscr{G}^+$ has a subgame-perfect ε-equilibrium in pure strategies for every $\varepsilon > 0$. Moreover, if all transitions in G are deterministic, then G has a subgame-perfect 0-equilibrium in pure strategies.*

The idea of the construction. Take an arbitrary game $G \in \mathscr{G}^+$. We assume that, in all absorbing states s, player i_s has only one action. (Otherwise, we reduce the game by deleting all actions in state s except one that offers player i_s the highest payoff.) This will imply that, as soon as absorption takes place, play is over from a strategic point of view. For a real vector $\alpha = (\alpha_s)_{s\in S}$ and a starting state $s \in S$, let $V_s(\alpha)$ denote the set of joint pure strategies π with the following properties:

(1) π is absorbing, i.e., according to π, absorption occurs eventually with probability 1.

(2) For any history h ending in a state $t \in S$, the joint strategy π offers player i_t an expected continuation reward of at least α_t, i.e., $\phi^{i_t}(\pi[h]) \geq \alpha_t$.[3]

Let Λ denote the set of all vectors $\alpha = (\alpha_s)_{s\in S}$ such that $V_s(\alpha)$ is nonempty for all $s \in S$. For any $\alpha \in \Lambda$ and state $s \in S$, let

$$\delta_s(\alpha) := \max_{a\in A_s} \inf_{\pi_t \in V_t(\alpha),\ \forall\, t\in S} \sum_{t\in S} [p_s(a, t) \cdot \phi^{i_s}(\pi_t)].$$

Intuitively, $\delta_s(\alpha)$ is player i_s's best reward when he plays an action a in state s, and subsequently, starting from the next state t, all players (including player i_s) minimize player i_s's reward by playing some joint strategy $\pi_t \in V_t(\alpha)$. In fact, $\delta_s(\alpha)$ can be seen as some punishment level for player i_s in state s, given punishment can only be executed within the sets $\pi_t \in V_t(\alpha)$, $t \in S$. Let $\delta = (\delta_s)_{s\in S}$.

[3] In the formal proof, we will use a slightly different definition for the sets $V_s(\alpha)$.

Sufficient conditions for subgame-perfect 0-equilibria in $\mathscr{G}^+$. Assume that there exists an $\bar{\alpha} \in \Lambda$ such that

(1) $\bar{\alpha}$ is a fixed point of δ, i.e., $\delta(\bar{\alpha}) = \bar{\alpha}$;

(2) the infimum is attained in the definition of $\delta_s(\bar{\alpha})$ for all states $s \in S$, i.e., there are $\pi_t^s \in V_t(\bar{\alpha})$, $\forall\, t \in S$, such that

$$\delta_s(\bar{\alpha}) = \max_{a \in A_s} \sum_{t \in S} [p_s(a, t) \cdot \phi^{i_s}(\pi_t^s)].$$

Under these conditions, one can construct a subgame-perfect 0-equilibrium $\bar{\pi}$ as follows. Take an initial state $s \in S$ and an arbitrary game plan $\pi \in V_s(\bar{\alpha})$. When starting in state s, the joint strategy $\bar{\pi}$ prescribes to play according to $\pi^1 := \pi$, as long as all players follow the prescriptions of π^1. If, on the other hand, in some state t, the controlling player i_t deviates to another action, by which play moves to some state u, then $\bar{\pi}$ prescribes to switch to $\pi^2 := \pi_u^t$. Then, π^2 will be used as long as all players follow the prescriptions by π^2. And similarly, if, in a state w, player i_w deviates by which play moves to state z, then $\bar{\pi}$ prescribes to switch to $\pi^3 := \pi_z^w$, and so on.

The main point of the construction of $\bar{\pi}$ is the following. Suppose that the players are currently using π^1 and play is in state t. Because of property (2) for π^1, player i_t expects a reward of at least $\bar{\alpha}_t$ from π^1. On the other hand, if player i_t decides to deviate in state t, then, by property (2), the new π^2 offers player i_t a reward of at most $\delta_t(\bar{\alpha})$ in expectation. By property (1), $\delta_t(\bar{\alpha}) = \bar{\alpha}_t$, which implies that such a single deviation is not profitable.

Based on this observation, we now provide an intuitive argument why $\bar{\pi}$ is a subgame-perfect 0-equilibrium. Because play is similar in any subgame, we only argue that $\bar{\pi}$ is a 0-equilibrium. Suppose that some player i intends to deviate by choosing another strategy τ^i. Because player i cannot profit from preventing absorption completely (recall that all payoffs in the absorbing states are nonnegative), we can assume that $(\bar{\pi}^{-i}, \tau^i)$ is absorbing. Therefore, τ^i only deviates from $\bar{\pi}^i$ finitely many times, with probability 1. Because we already argued that a single deviation cannot be profitable, it will follow that τ^i does not yield a better reward to player i than $\bar{\pi}^i$.

Unfortunately, it is not clear whether or not all games in $\mathscr{G}^+$ satisfy properties (1) and (2) above. We will show, nevertheless, that every game $G \in \mathscr{G}^+$ with deterministic transitions does have these two properties. As a consequence, G admits a subgame-perfect 0-equilibrium.

The existence of subgame-perfect ε-equilibria in $\mathscr{G}^+$. Regarding subgame-perfect ε-equilibria, however, it is possible to make progress even without properties (1) and (2). We propose the following *iterative method*. Let $\alpha_s^0 = 0$ for all states $s \in S$.[4] Obviously, for any $s \in S$, the set $V_s(\alpha^0)$ consists of all joint pure strategies, which are absorbing. Thus $\alpha^0 \in \Lambda$. For a general k, given $\alpha^k \in \Lambda$, we let $\alpha^{k+1} = \delta(\alpha^k)$. We then prove that $\alpha^{k+1} \in \Lambda$. This sequence α^k in Λ turns out to be nondecreasing. Because the payoffs are bounded, we may conclude that the sequence α^k has a finite limit, which we denote by α^*. Although we can show that $\alpha^* \in \Lambda$, it is not clear whether $\delta(\alpha^*) = \alpha^*$ holds (cf. the second concluding remark). Nevertheless, by exploiting the fact that α^* is a limit point of a successive application of δ, we are able to construct a subgame-perfect ε-equilibrium. It has a similar structure as $\bar{\pi}$ above, only the underlying sequence $(\pi^m)_{m \in \mathbb{N}}$ has to be chosen in a more clever manner.

Related iterative methods. Beside the method in Vartiainen [23], which was discussed in the introduction, we also mention an iterative method applied in Maitra and Sudderth [8, 9]. Their iterative method was used to show the existence of the value in two-player zero-sum stochastic games. They consider some auxiliary games, called leavable games, in which player 1 can decide when to "stop." The terminal payoff corresponding to stopping is a variable. They define an inductive process in which these terminal payoffs are, in turn, determined by the value of another leavable game. They show that the obtained sequence of values is monotonic and convergent.

3. A reduction. In this section, we will show that, for our main results, it is sufficient to guarantee the existence of pure subgame-perfect ε-equilibria in a subclass of $\mathscr{G}^+$. Let $\widetilde{\mathscr{G}}^+$ denote the class of games in $\mathscr{G}^+$, which satisfy

(1) From any state, the players have a joint strategy such that absorption eventually occurs with probability 1. (One can show that it is sufficient to consider pure stationary strategies.)

(2) In any absorbing state s, player i_s has precisely one action.

(3) In any nonabsorbing state s, any action of player i_s either leads to nonabsorbing states only, or it leads to one absorbing state with probability 1. In this sense, we will speak of nonabsorbing and absorbing actions.

(4) In any nonabsorbing state s, player i_s has precisely one absorbing action.

[4] In the formal proofs, we choose other initial values.

Now, we will provide a natural transformation, which reduces an arbitrary game $G \in \mathscr{G}^+$ into a game $\tilde{G} \in \tilde{\mathscr{G}}^+$. So, take a game $G \in \mathscr{G}^+$. We obtain $\tilde{G}$ in four steps as follows:

Step 1 for property (1). If G does not satisfy this property, then there must exist a state s such that, irrespective of the strategies of the players, the probability that play ever absorbs is zero. This means the payoffs are zeros regardless of what happens. For this reason, we replace all such states by absorbing states in which the controlling player has one action and all payoffs are equal to zero.

Step 2 for property (2). In any absorbing state s, we delete all actions of player i_s except one that offers player i_s the highest payoff in state s.

Step 3 for property (3). Suppose that, in some nonabsorbing state s, some action a of player i_s does not satisfy this property. Then, action a leads to the set of absorbing states with a positive probability. Given absorption takes place through action a, let w^j denote the conditional expected payoff for player j in the set of absorbing states. More formally,

$$w^j = \frac{\sum_{u \in S,\ u \text{ is absorbing}} (p_s(a, u) \cdot r_u^j)}{\sum_{u \in S, u \text{ is absorbing}} p_s(a, u)},$$

where r_u^j denotes the unique payoff (cf. property (2)) for player j in absorbing state u. Then, (i) we add two new states: (ia) we add a new absorbing state s^* in which the controlling player has one action and the payoff for every player j is equal to w^j, (ib) we add a new nonabsorbing state s' with one action for the controlling player from which transition occurs to state s^* with probability 1, and (ii) we replace each transition through action a to an absorbing state by a transition to state s'. We apply the same transformation to all actions that violate property (3).

Step 4 for property (4). If in some nonabsorbing state s, player i_s has more than one absorbing action, then we delete all absorbing actions of player i_s, except one that offers player i_s the highest payoff at absorption. We apply the same transformation to all such states. Suppose now, on the other hand, that there is a nonabsorbing state s in which player i_s has no absorbing action. Let S^0 denote the set of these states. Then, (i) we raise all payoffs in the absorbing states by 1, (ii) we add a new absorbing state t^* in which the controlling player has one action and all payoffs are equal to 0, (iii) in each state in S^0, we add a new absorbing action leading to state t^*.

It is clear that the new game $\tilde{G}$ satisfies all four properties, thus $\tilde{G} \in \tilde{\mathscr{G}}^+$.

Lemma 1. *Take an arbitrary game $G \in \mathscr{G}^+$, and transform G, according to the rules above, into a game $\tilde{G} \in \tilde{\mathscr{G}}^+$. If a subgame-perfect ε-equilibrium exists in $\tilde{G}$ in pure strategies for all $\varepsilon > 0$, then one also exists in G. Moreover, if a subgame-perfect 0-equilibrium exists in $\tilde{G}$ in pure strategies, then one also exists in G.*

Proof. We only show the second part, as the proof for the first part on subgame-perfect ε-equilibria is almost identical. Now, let π denote a pure subgame-perfect 0-equilibrium in the game $\tilde{G}$. We will show that π induces a pure subgame-perfect 0-equilibrium for the original game G.

We will only argue for the transformation in Step 4, as it is obvious for the other three steps. It is enough to show that π never prescribes to choose the new absorbing actions in states in S^0. Suppose $S^0 \neq \varnothing$, otherwise the statement is obvious. Let T denote the set of states after the transformation in Step 3 (T can differ from the original state space S because of Step 3). Because of the transformation in Step 1, $T - S^0 \neq \varnothing$. Consider first a state $s \in T - S^0$. In this state, player i_s can obtain a payoff of at least 1 by playing his absorbing action. Hence, his continuation reward from state s with regard to π is always at least 1. Notice that the payoffs in any absorbing state are either all at least 1 or all equal to 0. Hence, each player's (expected) continuation reward from state s with regard to π is always at least some $\rho_s > 0$. Because of the transformation in Step 1, there exists a state $t \in S^0$ with an action a such that action a leads to $S - S^0$ with a positive probability; say, q_a. Then, if player i_t plays action a in state t, his continuation reward with respect to π is at least $q_a \cdot \min_{s \in S - S^0} \rho_s > 0$. Hence the continuation reward of player i_t from state t with regard to π, and consequently, for all other players as well, is at least some $\rho_t > 0$. This means, in particular, that player i_t never plays his new absorbing action in state t. If $S^0 - \{t\} = \varnothing$, then we are done. On the other hand, if $S^0 - \{t\} \neq \varnothing$, then we can continue as before, since there exists again a state $u \in S^0 - \{t\}$ with an action b such that action b leads to $(T - S^0) \cup \{t\}$ with a positive probability. Since the number of states is finite, we obtain in finitely many steps a $\rho > 0$ such that the continuation rewards of the players with regard to π are always at least ρ. Hence the new absorbing actions in states in S^0 are never chosen, as desired. □

Because of the above lemma, we may restrict our investigation to games in the class $\tilde{\mathscr{G}}^+$.

An equivalent interpretation of games in $\widetilde{\mathscr{G}}^+$. Note that, in every state s of a game belonging to $\widetilde{\mathscr{G}}^+$, if player i_s plays his absorbing action, then play moves with probability 1 to the absorbing state associated with s; say, $t(s)$. Play is over from a strategic point of view once $t(s)$ is reached, and each player j receives an (expected) reward equal to $r^j_{t(s)}$. From now on, we will use the following equivalent interpretation of games in $\widetilde{\mathscr{G}}^+$: in every nonabsorbing state $s \in S$, if player i_s plays his absorbing action, then play terminates, with *payoff* $r^j_s := r^j_{t(s)}$ to every player j. If the players never use their absorbing actions, play continues forever with payoff 0 to every player. In this interpretation, the absorbing states play no role. Playing the absorbing action in a nonabsorbing state will be called *quitting*. Moreover, the absorbing actions will be called *quitting actions*, and all other actions will be called *nonquitting actions*.

4. Game plans and viability. Take a stochastic game $G \in \widetilde{\mathscr{G}}^+$ and an initial state $s \in S$. For a joint pure strategy $\pi \in \Pi_s$, let H^π denote the set of those histories, which have a positive probability with respect to π. A function g from histories to actions is called a *complete game plan* for initial state s if g equals the restriction $\pi_{|H^\pi}$ of some joint pure strategy $\pi \in \Pi_s$ to the set H^π. Let $H^g = H^\pi$ denote the domain of g. Thus $g(h) = \pi(h)$ for all $h \in H^g$. Clearly, different joint pure strategies can induce the same complete game plan. The idea of a complete game plan is that it provides a prescription for the choice of an action during the whole play, if no player ever deviates from the plan. We say that a game plan g is stationary if $g(h)$ only depends on the final state of h.

It is also possible to construct a complete game plan g for initial state s inductively as follows:

(1) At stage 1, an action $g(s) \in A_s$ is chosen. Define $H^{g,1} = \{(s)\}$.

(2) At stage $m+1$ $(m \geq 1)$, an action $g(h) \in A_u$ is chosen for all histories h of the form $h = h^m \oplus (t, g(h^m), u)$, where $h^m \in H^{g,m}$, where t denotes the final state of h^m, and where $u \in S_t(g(h^m))$. Let $H^{g,m+1}$ denote the set of such histories h.

The domain of such a constructed game plan g is given by $H^g = \bigcup_{m\in\mathbb{N}} H^{g,m}$.

A function g from histories to actions is called a *truncated game plan* for initial state s if g equals the restriction $\pi_{|W}$ of a joint pure strategy π to a set $W \subsetneqq H^\pi$ of histories such that W satisfies: if $h \in W$ is an arbitrary history; say up to stage m, then, for any stage $l < m$, the part of h up to stage l also belongs to W. An equivalent formulation of this property of the set W is that if a history does not belong to W, then neither does any extension of this history up to larger stages. Note that if the players follow the prescriptions of a truncated game plan g, then it may happen that a history $h \in H^\pi - W$ occurs. In this case, g provides no further prescriptions, and we will say that g *expires*. Also, for the truncated game plan g, let $H^g = W$ denote the domain of g.

Example 1. Consider a game G with three players and three states s, t, and u. In state s, player 1 can either play a nonquitting action a_s leading to state s with probability $\frac{1}{2}$ and to state t with probability $\frac{1}{2}$, or quit. In state t, player 2 can either play a nonquitting action a_t leading to state u with probability 1, or play the quitting action. In state u, player 3 can only quit. Consider state s as the initial state. An example of a complete game plan is $\hat{g}$, which prescribes action a_s in state s until play arrives at state t and then quits in state t. More formally, $H^{\hat{g}}$ consists of all histories that are either of the form $(s, a_s, s, a_s, \ldots, s, a_s, s)$ or of the form $(s, a_s, s, a_s, \ldots, s, a_s, t)$, and $\hat{g}$ assigns action a_s to the former ones and quitting to the latter ones. An example of a truncated game plan is g', which prescribes action a_s in state s until play arrives at state t. More formally, $H^{g'}$ consists of all histories of the form $(s, a_s, s, a_s, \ldots, s, a_s, s)$, and g' assigns action a_s to all these histories.

Consider a (complete or truncated) game plan g for initial state s. For any $h \in H^g$, we can define the *continuation game plan* $g[h]$ of g with respect to h by $g[h](h') = g(h \oplus h')$ for all h' with $h \oplus h' \in H^g$, just as in the case of strategies. Notice that $g[h]$ is a game plan for the final state of h. In the following, we will sometimes use the notation H^g_t to denote the histories in H^g with final state t. Thus, if $h \in H^g_t$, then $g[h]$ is a game plan for state t. Note that $g[h]$ is complete if g is complete.

A complete game plan g is called *quitting*, if playing according to g eventually leads to quitting, with probability 1. For a quitting game plan g, the expected payoff to a player i is denoted by $\phi^i(g)$. Note that any continuation game plan $g[h]$ of a quitting game plan g is also quitting.

Take a real vector $\alpha = (\alpha_t)_{t\in S}$. A quitting game plan g is called *viable* with respect to α if $\phi^{i_t}(g[h]) \geq \alpha_t$ holds for all $t \in S$ and all $h \in H^g_t$. This means that, whenever play is in some state t, and play is according to g, the controlling player i_t can expect a payoff of at least α_t. In particular, since g always prescribes one action with probability 1, viability of g with respect to α implies that termination can only take place at states t with the property $r^{i_t}_t \geq \alpha_t$. A state t with the property $r^{i_t}_t \geq \alpha_t$ is called a *quitting state* with respect to α.

Let $V_s(\alpha)$ denote the set of game plans for initial state s that are viable with respect to α. A specific viable game plan will be denoted by v. Note that, for a game-plan $v \in V_s(\alpha)$ and a history $h \in H^v$ with final state t, the continuation game plan $v[h]$ is a viable game plan in $V_t(\alpha)$. Notice that viable game plans are, by definition, always quitting and complete.

Example 2. Consider the game G from Example 1 and game plan $\hat{g}$ for initial state s as defined in that example. Let the payoff vectors be given by $(2,2,2)$, $(3,1,2)$, and $(1,2,2)$ for s, t, and u, respectively. Let $\alpha_s = \alpha_u = 2$ and $\alpha_t = 1$. According to $\hat{g}$, whenever play is in state s, player 1's expected payoff is equal to $3 > 2 = \alpha_s$, and whenever play is in state t, player 2 receives $1 = \alpha_t$. Hence $\hat{g}$ is viable with respect to α.

Consider two states t, $u \in S$, and a real vector $\alpha = (\alpha_t)_{t \in S}$. We define

$$\beta_t(u, \alpha) = \begin{cases} \inf\limits_{v \in V_u(\alpha)} \phi^{i_t}(v) & \text{when } V_u(\alpha) \neq \varnothing \\ \infty & \text{otherwise,} \end{cases}$$

which is the highest lower bound for the payoff to player i_t with regard to viable game plans starting in state u. We do not know if the infimum is always attained.

Now, we can also define such a bound for an action $a \in A_t$ of player i_t in state t as follows:

$$\gamma_t(a, \alpha) = \begin{cases} \sum\limits_{u \in S} p_t(a, u)\beta_t(u, \alpha) & \text{if action } a \text{ is nonquitting} \\ r_t^{i_t} & \text{if action } a \text{ is quitting.} \end{cases}$$

Furthermore, let

$$\delta_t(\alpha) = \max_{a \in A_t} \gamma_t(a, \alpha).$$

Recall from §2 that $\delta_t(\alpha)$ can be seen as a punishment level for player i_t in state t, given punishment can only be executed with game plans that are viable for α.

We finally define $B_t(\alpha)$ as the set of those actions a for player i_t in state t for which $\gamma_t(a, \alpha) = \delta_t(\alpha)$.

By the definition of $\delta_t(\alpha)$, the set $B_t(\alpha)$ is always nonempty. Observe that if $\delta_t(\alpha) = \infty$, then $B_t(\alpha)$ consists of those nonquitting actions a in state t for which there is a state $u \in S_t(a)$ with $V_u(\alpha) = \varnothing$.

The following lemma provides useful properties of these functions. Its proof follows straightforwardly from the definitions.

Lemma 2. *If $\alpha \geq \bar{\alpha}$, then we have for all states $t \in S$ that*
(1) $V_t(\alpha) \subseteq V_t(\bar{\alpha})$;
(2) $\beta_t(u, \alpha) \geq \beta_t(u, \bar{\alpha})$ *for all* $u \in S$;
(3) $\gamma_t(a, \alpha) \geq \gamma_t(a, \bar{\alpha})$ *for all* $a \in A_t$;
(4) $\delta_t(\alpha) \geq \delta_t(\bar{\alpha})$.

5. Subgame-perfect ε-equilibria in pure strategies in $\widetilde{\mathscr{G}}^+$. In this section, we prove the existence of subgame-perfect ε-equilibria, in pure strategies, for all games in the class $\widetilde{\mathscr{G}}^+$. So, consider an arbitrary game $G \in \widetilde{\mathscr{G}}^+$. We start with an *iterative scheme* (cf. §2), which will appear to converge to a finite limit. We define a sequence $\alpha^0, \alpha^1, \ldots$ of vectors, where $\alpha^k = (\alpha_t^k)_{t \in S} \in \mathbb{R}^{|S|}$ by

$$\alpha_t^k = \begin{cases} r_t^{i_t} & \text{if } k = 0 \\ \delta_t(\alpha^{k-1}) & \text{if } k > 0 \end{cases}$$

for all states $t \in S$.

Lemma 3. *For all $k \geq 0$, it holds that $\alpha^{k+1} \geq \alpha^k$.*

Proof. The proof is by induction. Let a_t denote the quitting action in state t. For $k = 0$, we have for any state $t \in S$ that

$$\alpha_t^1 = \delta_t(\alpha^0) \geq \gamma_t(a_t, \alpha^0) = r_t^{i_t} = \alpha_t^0.$$

Assume now that $\alpha^k \geq \alpha^{k-1}$ holds for some $k \geq 1$. Then, by Lemma 2–4, we obtain for any state $t \in S$ that

$$\alpha_t^{k+1} = \delta_t(\alpha^k) \geq \delta_t(\alpha^{k-1}) = \alpha_t^k,$$

which completes the proof. $\square$

Observe the following. For any state $t \in S$, at any iteration level k, either $\alpha_t^k = \infty$ or α_t^k is bounded from above by the maximal payoff of the game. Since the sequence α_t^k $(k = 0, 1, \dots)$ is nondecreasing, α_t^k will either converge to a finite limit, or, after a finite number of iterations, $\alpha_t^k = \infty$. We will call this limit α_t^*, and set $\alpha^* = (\alpha_t^*)_{t \in S}$.

Example 3. Consider the following game G. There are four players, so $N = \{1, 2, 3, 4\}$. The state space is $S = \{s_1, s_2, s_3, s_4\}$ with player i controlling state s_i. In state s_i, player i's action space is $A_{s_i} = \{a_1^i, a_2^i, a_3^i, a_4^i\}$, where action a_j^i with $j \neq i$, leads to state s_j with probability 1, whereas action a_i^i is quitting. The payoffs for quitting are

$$r_{s_1} = (1, 1, 3, 1), \quad r_{s_2} = (1, 1, 1, 1), \quad r_{s_3} = (1, 2, 2, 1), \quad r_{s_4} = (2, 1, 1, 2).$$

Obviously, $G \in \tilde{\mathscr{G}}^+$ and all transitions in G are deterministic. We are going to perform the iteration scheme defined above, and we will see that it allows us to construct a pure subgame-perfect 0-equilibrium. For simplicity, a quitting game plan for initial state $t_1 \in S$ will be denoted by $(t_1, t_2, \dots, t_m, *)$, where $t_l \neq t_{l+1}$ for all $l = 1, \dots, m-1$, with the interpretation that in state t_1 action $a_{t_2}^{t_1}$ should be played, leading to state t_2, then in state t_2 action $a_{t_3}^{t_2}$ should be played, leading to state t_3, and so on until quitting should take place in state t_m.

Step 0. Initially, $\alpha_{s_1}^0 = \alpha_{s_2}^0 = 1$ and $\alpha_{s_3}^0 = \alpha_{s_4}^0 = 2$. Notice that, since player 4 receives less than $\alpha_{s_4}^0 = 2$ in all states except his own state s_4, all game plans, which are viable with regard to α^0 and which start in state s_4, will eventually quit in state s_4.

Regarding player 1: Because of the above observation, we obtain $\beta_{s_1}(s_4, \alpha^0) = 2$ and $a_4^1 \in B_{s_1}(\alpha^0)$. Hence $\alpha_{s_1}^1 = 2$.

Regarding player 2: We have $\beta_{s_2}(s_1, \alpha^0) = 1$ as the game plan $(s_1, s_4, *)$ (or simply quitting in state s_1 immediately) is viable with regard to α^0. Also, $\beta_{s_2}(s_3, \alpha^0) = 1$, as $(s_3, s_1, *)$ is viable with regard to α^0. Clearly, $\beta_{s_2}(s_4, \alpha^0) = 1$. Thus $\beta_{s_2}(t, \alpha^0) = 1$ for all states $t \in S$, yielding $\alpha_{s_2}^1 = 1$.

Regarding player 3: Since $(s_1, s_2, *)$ is viable with respect to α^0, it follows that $\alpha_{s_3}^1 = 2$.

Regarding player 4: Of course, $\alpha_{s_4}^1 = 2$.

Step 1. We obtained $\alpha_{s_1}^1 = \alpha_{s_3}^1 = \alpha_{s_4}^1 = 2$ and $\alpha_{s_2}^1 = 1$. The main difference is that quitting in state s_1 is no longer viable for player 1. This means that $(s_3, s_1, *)$ is not viable with regard to α^1. Hence $\beta_{s_2}(s_3, \alpha^1) = 2$ yielding $\alpha_{s_2}^2 = 2$. Thus $\alpha_t^2 = 2$ for all states $t \in S$.

Step 2. We obtained $\alpha_t^2 = 2$ for all states $t \in S$. It is easy to see that $\alpha^k = \alpha^2$ for all $k > 2$, implying $\alpha_t^* = 2$ for all states $t \in S$.

Consider the following game plans, which are all viable with respect to α^*:

$$g_{s_1} = (s_1, s_4, *), \quad g_{s_2} = (s_2, s_3, *), \quad g_{s_3} = (s_3, *), \quad g_{s_4} = (s_4, *).$$

Now, we obtain a pure subgame-perfect 0-equilibrium as follows. From any initial state $s \in S$, game plan g_s should be played. If any player along the way deviates from g_s to a nonquitting action, by which play moves to state t, then game plan g_t should be played from state t. And similarly, if a deviation occurs from g_t to a nonquitting action leading to state u, then game plan g_u should be played from state u, and so on. It is easy to check that this prescription provides a subgame-perfect 0-equilibrium.

The first part of the proof for the main result is to show that viable game plans exist for any starting state with respect to α^k for all k (cf. part 1 of Lemma 8). We start with the following lemma, which states that the concatenation of a best action for α^{k-1} or α^k with a viable game plan for α^k is also a viable game plan for α^k.

Lemma 4. *Suppose $V_t(\alpha^k) \neq \varnothing$ and let $v_t \in V_t(\alpha^k)$ for all $t \in S$. Choose an initial state $s \in S$, and an action $a \in A_s$. Construct a game plan v for initial state s as follows*:

(1) *At stage* 1, *player i_s plays action a*;

(2) *If action a is nonquitting and play reaches state $t \in S_s(a)$, then from stage* 2 *onwards, game plan v_t will be played* (*with forgetting the history before stage* 2 *and considering state t as the initial state*).

If either $a \in B_s(\alpha^{k-1})$ or $a \in B_s(\alpha^k)$, then $v \in V_s(\alpha^k)$.

Proof. First assume that a is the quitting action. If $a \in B_s(\alpha^{k-1})$, then $\delta_s(\alpha^{k-1}) = \gamma_s(a, \alpha^{k-1})$; hence, we obtain

$$\alpha_s^k = \delta_s(\alpha^{k-1}) = \gamma_s(a, \alpha^{k-1}) = r_s^{i_s}.$$

If $a \in B_s(\alpha^k)$, then $\delta_s(\alpha^k) = \gamma_s(a, \alpha^k)$, hence

$$\alpha_s^k \leq \alpha_s^{k+1} = \delta_s(\alpha^k) = \gamma_s(a, \alpha^k) = r_s^{i_s},$$

where the inequality follows from Lemma 3. In either case, quitting is viable with respect to α^k.

Now, assume that a is nonquitting. Since $v_t \in V_t(\alpha^k)$ for every state $t \in S_s(a)$, we conclude that the game plan v is quitting and that $\phi^{i_u}(v[h]) \geq \alpha_u^k$ holds for any $u \in S$ and any history $h \in H_u^v$ that reaches stage 2. It remains to show that $\phi^{i_s}(v) \geq \alpha_s^k$. As

$$\phi^{i_s}(v) = \sum_{t \in S_s(a)} p_s(a,t)\phi^{i_s}(v_t) \geq \sum_{t \in S_s(a)} p_s(a,t)\beta_s(t,\alpha^k) = \gamma_s(a,\alpha^k),$$

it suffices to verify $\gamma_s(a,\alpha^k) \geq \alpha_s^k$. We distinguish two cases: $a \in B_s(\alpha^{k-1})$ or $a \in B_s(\alpha^k)$. If $a \in B_s(\alpha^{k-1})$, we obtain

$$\alpha_s^k = \delta_s(\alpha^{k-1}) = \gamma_s(a,\alpha^{k-1}) \leq \gamma_s(a,\alpha^k),$$

where the inequality follows from Lemmas 2–3 and 3. On the other hand, if $a \in B_s(\alpha^k)$, we have

$$\alpha_s^k \leq \alpha_s^{k+1} = \delta_s(\alpha^k) = \gamma_s(a,\alpha^k),$$

where the inequality follows from Lemma 3. Hence $\phi^{i_s}(v) \geq \alpha_s^k$ as desired. ☐

The next lemma, which is a generalization of Lemma 4, states that if players start by playing best actions for α^{k-1} or α^k, and eventually switch to playing a viable game plan for α^k, then this concatenation is a viable game plan for α^k.

Lemma 5. *Suppose $V_t(\alpha^k) \neq \varnothing$ and let $v_t \in V_t(\alpha^k)$ for all $t \in S$. Choose an initial state $s \in S$, and let g be a truncated game plan for s that only uses actions from the sets $B_u(\alpha^{k-1})$ and $B_u(\alpha^k)$ for every state $u \in S$, and that expires with probability* 1. *Let v denote the complete game plan for initial state s, according to which*

(1) *from stage* 1 *onward, game plan g is executed, and*

(2) *when g expires; say, in state t, game plan v_t will be played (with forgetting the history induced by g and considering state t as the initial state).*

Then $v \in V_s(\alpha^k)$.

Proof. It is clear that v is a quitting game plan. We need to prove that v is viable with respect to α^k. For any $m \in \mathbb{N}$, consider the following game plan v_m, by adapting v: if, at stage m, the game plan g has not expired yet, and play is in some state t, then start game plan v_t. Note that v_m is viable with respect to α^k for all m, which follows by repeated application of Lemma 4. Let $\varepsilon > 0$. Since g expires with probability 1, we can choose m so large that the probability that g expires before stage m is so close to 1 that

$$\phi^{i_s}(v_m) \leq \phi^{i_s}(v) + \varepsilon.$$

Because $v_m \in V_s(\alpha^k)$, we have

$$\phi^{i_s}(v_m) \geq \alpha_s^k.$$

Therefore

$$\phi^{i_s}(v) \geq \phi^{i_s}(v_m) - \varepsilon \geq \alpha_s^k - \varepsilon.$$

Because $\varepsilon > 0$ was arbitrary, we proved $\phi^{i_s}(v) \geq \alpha_s^k$.

Observe that, for any $t \in S$ and any history $h \in H_t^v$, the above reasoning can also be given for $v[h]$ to show that $\phi^{i_t}([h]) \geq \alpha_t^k$. Hence $v \in V_s(\alpha^k)$. ☐

Let $k \in \mathbb{N}$. For two states s and t, we write $s \preceq^k t$ if state t can eventually be visited with a positive probability when starting in s, by only using actions in the sets $B_u(\alpha^k)$, $u \in S$. This relation $\preceq^k$ is obviously transitive. With respect to $\preceq^k$, a nonempty set $Q \subseteq S$ is called closed, if for every $s \in Q$ there is no $t \in S - Q$ such that $s \preceq^k t$. A closed set $Q \subseteq S$ is called minimal closed if Q contains no proper subset, which is closed. We will call every minimal closed set a *persistent set* and its elements persistent states with respect to α^k. It is clear that there always exists a persistent set. Thus we have the following properties for persistent states: (1) from any nonpersistent state, we can eventually reach the set of persistent states with probability 1, by only using actions in the sets $B_u(\alpha^k)$, $u \in S$, (2) a persistent set P^k cannot be left through actions in the sets $B_u(\alpha^k)$, $u \in P^k$, and (3) if s and $t \neq s$ belong to the same persistent set P^k, then t can eventually be visited when starting in s with probability 1, by only using actions in the sets $B_u(\alpha^k)$, $u \in P^k$.

The following two lemmas demonstrate useful properties of persistent states, which are still needed to show that viable game plans exist for any starting state with respect to α^k for all k (cf. part 1 of Lemma 8). Assume that viable game plans exist for every state at a given iteration step. The next lemma states that a quitting action in a persistent state is always a best action for α^0. Moreover, for a persistent state, best actions from the previous iteration step are still best actions.

Lemma 6. *Suppose that $V_u(\alpha^k) \neq \varnothing$ for all $u \in S$. Take a persistent state $s \in S$ with respect to α^k. Then*
(1) *For $k=0$: If a is quitting, then $a \in B_s(\alpha^0)$.*
(2) *For $k>0$: If $a \in B_s(\alpha^{k-1})$, then $a \in B_s(\alpha^k)$.*

Proof. Let $a \in B_s(\alpha^{k-1})$ if $k>0$ or let a be quitting if $k=0$. We will show that $a \in B_s(\alpha^k)$. Let P^k denote the persistent set with respect to α^k that state s belongs to. For every state $t \in P^k$, we define a game plan v_t, starting in state t, according to which

(1) From stage 1 onward, starting in state t, a truncated game plan will be used to eventually visit state s, by only using actions in the sets $B_u(\alpha^k)$, $u \in P^k$. Such a game plan exists by property (3) of persistent states. (If $t=s$, then this truncated game plan is empty.) Notice that the game plan expires in state s with probability 1.

(2) When s is reached; say, at stage $m \geq 1$, action a will be played.

(3) If a is nonquitting, choose viable game plans $w_u \in V_u(\alpha^k)$ for all $u \in S_s(a)$. From stage $m+1$ onward, if $u' \in S_s(a)$ denotes the state at stage $m+1$, the viable game plan $w_{u'}$ will be played (with forgetting the history before stage $m+1$ and considering state u' as the initial state).

We will now show that $v_t \in V_t(\alpha^k)$ for all $t \in P^k$. We distinguish between two cases.

If $k>0$, then $a \in B_s(\alpha^{k-1})$. If a is quitting, then $\alpha_s^k = \delta_s(\alpha^{k-1}) = \gamma_s(a, \alpha^{k-1}) = r_s^{i_s}$, hence quitting is viable with respect to α^k. Therefore $v_t \in V_t(\alpha^k)$ because of Lemma 5. If a is nonquitting, then $v_t \in V_t(\alpha^k)$ follows directly from Lemma 5.

If $k=0$, then a is quitting, and as quitting is always viable with respect to α^0, viability of v_t follows by Lemma 5.

We claim that

$$\beta_s(t, \alpha^k) \leq \gamma_s(a, \alpha^k)$$

for all $t \in P^k$. As $v_t \in V_t(\alpha^k)$ for all $t \in P^k$, we have $\beta_s(t, \alpha^k) \leq \phi^{i_s}(v_t)$. If a is quitting, then each game plan v_t terminates at s with payoff $r_s^{i_s}$ for i_s, in which case the claim follows from:

$$\beta_s(t, \alpha^k) \leq \phi^{i_s}(v_t) = r_s^{i_s} = \gamma_s(a, \alpha^k).$$

If a is nonquitting, then each game plan v_t induces the payoff

$$\phi^{i_s}(v_t) = \sum_{u \in S_s(a)} p_s(a,u) \phi^{i_s}(w_u).$$

The claim now follows from

$$\beta_s(t, \alpha^k) \leq \sum_{u \in S_s(a)} p_s(a,u) \Big[\inf_{w_u \in V_u(\alpha^k)} \phi^{i_s}(w_u) \Big] = \sum_{u \in S_s(a)} p_s(a,u) \beta_s(u, \alpha^k) = \gamma_s(a, \alpha^k),$$

because the choice of w_u was arbitrary in $V_u(\alpha^k)$ for all $u \in S_s(a)$.

Now, take $b \in B_s(\alpha^k)$. We claim that

$$\gamma_s(b, \alpha^k) \leq \gamma_s(a, \alpha^k).$$

If b is quitting, this claim follows from $\gamma_s(b, \alpha^k) = r_s^{i_s} \leq \gamma_s(a, \alpha^k)$. If b is nonquitting, then, because of $s \in P^k$ and $b \in B_s(\alpha^k)$, we have $S_s(b) \subseteq P^k$. Therefore the claim follows from:

$$\gamma_s(b, \alpha^k) = \sum_{t \in S_s(b)} p_s(b,t) \beta_s(t, \alpha^k) \leq \sum_{t \in S_s(b)} p_s(b,t) \gamma_s(a, \alpha^k) = \gamma_s(a, \alpha^k).$$

Because $b \in B_s(\alpha^k)$, we may derive

$$\delta_s(\alpha^k) = \gamma_s(b, \alpha^k) \leq \gamma_s(a, \alpha^k),$$

which implies $a \in B_s(\alpha^k)$ as desired. □

Lemma 7. *If t is a quitting state with respect to α^k and if t is persistent with respect to α^k, then t is a quitting state with respect to α^{k+1}.*

Proof. Let t be a persistent quitting state with respect to α^k and let a denote the quitting action. If $k=0$, then $a \in B_t(\alpha^k)$ by Lemma 6. If $k>0$, then $a \in B_t(\alpha^{k-1})$, since $\delta_t(\alpha^{k-1}) = \alpha_t^k = r_t^{i_t} = \gamma_t(a, \alpha^{k-1})$. So if $k>0$, $a \in B_t(\alpha^k)$ also follows by Lemma 6. Consequently, $\alpha_t^{k+1} = \delta_t(\alpha^k) = \gamma_t(a, \alpha^k) = r_t^{i_t}$. □

We are now ready to prove that viable game plans exist for any starting state with respect to α^k for all k, and consequently, that the iterative scheme converges to a finite limit.

Lemma 8. (1) *$V_s(\alpha^k) \neq \varnothing$ for all $k \geq 0$ and all $s \in S$,*
(2) *Every persistent set with respect to α^k contains quitting states with respect to α^k.*
(3) *The limit α^* of the iterative process α^k is finite.*

Proof. We prove 1 and 2 by induction on k. Notice that both 1 and 2 are trivially true for $k = 0$. Now, assume 1 and 2 are true for some $k \geq 0$.

To prove claim 1 for $k+1$, notice that a truncated game plan for $s \in S$ exists that only uses actions in the sets $B_u(\alpha^k)$ ($u \in S$), and that expires with probability 1 in a persistent set with respect to α^k. This follows by property (1) of persistent states. Moreover, the truncated game plan can be extended to a truncated game plan that expires with probability 1 at a quitting state with respect to α^k, only by actions in the sets $B_u(\alpha^k)$, by property (3) of persistent sets and by the assumption that 2 holds for k. We then complete the game plan, which we denote by w_s, by choosing the quitting action when a quitting state with respect to α^k is reached. By Lemma 7, the state is also quitting with respect to α^{k+1}, and we may apply Lemma 5, to derive that $w_s \in V_s(\alpha^{k+1})$.

To prove 2 for $k+1$, let P^{k+1} be a persistent set with respect to α^{k+1}, choose $s \in P^{k+1}$ arbitrarily, and construct the game plan $w_s \in V_s(\alpha^{k+1})$ as above. We claim that the game plan w_s only visits states in P^{k+1}. This is trivial for state s visited at stage 1. Assume it is true for a state t visited at stage m. By construction of w_s, the action at state t is chosen from $B_t(\alpha^k)$. By Lemma 6, this action is also in $B_t(\alpha^{k+1})$. Therefore the state visited at stage $m+1$ is again in P^{k+1}, by property (2) of persistent sets. This demonstrates that w_s terminates with probability 1 in P^{k+1}. Because $w_s \in V_s(\alpha^{k+1})$, it follows that P^{k+1} contains quitting states with respect to α^{k+1}.

To prove 3, notice that because of 1, each α_s^k is finite for all $s \in S$ and $k \in \mathbb{N}$. Moreover, as each α_s^k is a convex combination of payoffs in the game, it is bounded from above by the maximal payoff. Since the sequence $(\alpha^k)_{k\in\mathbb{N}}$ is nondecreasing, it converges to a finite limit. □

Now, we are ready to show the Main Theorem restricted to games in $\tilde{\mathscr{G}}^+$.

Theorem 9. *In every stochastic game G in class $\tilde{\mathscr{G}}^+$, there exists a subgame-perfect ε-equilibrium in pure strategies for every $\varepsilon > 0$. Moreover, if all transitions in G are deterministic, then G has a subgame-perfect 0-equilibrium in pure strategies.*

Proof (General Transitions). We start by showing the case of general transitions. Take a stochastic game G in $\tilde{\mathscr{G}}^+$. We assume that all payoffs at quitting are at least 1; otherwise we can raise all payoffs at quitting by 1 (for any $\varepsilon \geq 0$, any subgame-perfect ε-equilibrium in this modified game is also a subgame-perfect ε-equilibrium in the original game). Take an initial state $s \in S$. Let $\varepsilon > 0$ and let $k \in \mathbb{N}$ be so large that

$$\|\alpha^* - \alpha^k\| \leq \frac{\varepsilon}{4|S|},$$

where the norm is the maximum norm, and where $|S|$ equals the number of states.

We will now define a joint pure strategy π^ε, and show that π^ε is a subgame-perfect ε-equilibrium for initial state s.

Step 1. Definition of π^ε. We will define the joint pure strategy π^ε inductively. Let $s^1 = s$ and take an arbitrary game plan $v^1 \in V_{s^1}(\alpha^k)$. When starting in state s^1, the joint strategy π^ε prescribes to play according to the game plan v^1, as long as all players follow the prescriptions of v^1. If, on the other hand, in some state, the controlling player ignores the prescription by v^1 and deviates to a nonquitting action, by which play moves to some state s^2, then π^ε prescribes to switch to a certain new game plan $v^2 \in V_{s^2}(\alpha^{k+1})$. This game plan v^2, to be specified later, will be used as long as all players follow the prescriptions by v^2. Similarly, if deviation occurs at some point to a nonquitting action, then π^ε prescribes another new game plan $v^3 \in V_{s^3}(\alpha^{k+2})$ from the state s^3 right after the deviation, and so on. Thus, with respect to π^ε, a game plan is active at any point during play.

We will now describe the choice of these game plans after a deviation takes place. Suppose the players are expected to use game plan v^m, but in state t, player i_t deviates to nonquitting action a. Let s^{m+1} denote the state to which transition occurs through action a. Then, the new game plan v^{m+1} is chosen such that $v^{m+1} \in V_{s^{m+1}}(\alpha^{k+m})$ and the expected reward satisfies

$$\phi^{i_t}(v^{m+1}) \leq \beta_t(s^{m+1}, \alpha^{k+m}) + \frac{\varepsilon}{2^{m+1}}.$$

Such a game plan exists by Lemma 8 and by the definition of the function β.

Step 2. π^ε is a subgame-perfect ε-equilibrium for initial state s. To prove this, we will show that π^ε is an ε-equilibrium for initial state s. Since the structure of any continuation strategy $\pi^\varepsilon[h]$ is almost identical to that of π^ε (the only difference is that $\pi^\varepsilon[h]$ starts with a continuation game plan of v^m for some m), a similar proof can be given that $\pi^\varepsilon[h]$ is an ε-equilibrium in the subgame after an arbitrary history h.

Take a player i, and a pure strategy $\tilde{\sigma}^i$ for player i. We will show that player i cannot improve his expected payoff by more than ε if he deviates from the strategy $\pi^{\varepsilon,i}$ to $\tilde{\sigma}^i$, i.e.,

$$\phi_s^i(\tilde{\sigma}^i, \pi^{\varepsilon,-i}) \le \phi_s^i(\pi^{\varepsilon,i}, \pi^{\varepsilon,-i}) + \varepsilon. \tag{1}$$

Note that it suffices to only consider pure deviations for player i, as for every $\delta > 0$, every player has a pure δ-best response to any joint strategy of his opponents. This follows from the fact that against fixed strategies of his opponents, every strategy for player i is equivalent with a mixed strategy, i.e., with a convex combination of pure strategies (cf. Aumann [1]); cf. also Theorem 1 in Monash [13].

Let σ^i be the strategy for player i, which follows the prescriptions of $\tilde{\sigma}^i$ until, during play, a history h occurs, with a final state u controlled by player i such that either

(1) the probability that $(\tilde{\sigma}^i[h], \pi^{\varepsilon,-i}[h])$ ever prescribes quitting is less than $1/\bar{r}$, where $\bar{r}$ is the maximal payoff in the game, or

(2) $\tilde{\sigma}^i[h]$ prescribes to quit at u.

In both cases, $\sigma^i[h]$ tells player i to play according to $\pi^{i,\varepsilon}[h]$. Notice that the expected payoff for player i with regard to $(\tilde{\sigma}^i[h], \pi^{\varepsilon,-i}[h])$ is at most his expected payoff with regard to $(\sigma^i[h], \pi^{\varepsilon,-i}[h])$. In case $\tilde{\sigma}^i[h]$ prescribes quitting, this follows from the fact that $\pi^\varepsilon[h]$ is viable with respect to α^*, and since $\alpha_u^* \ge r_u^i$. In case the probability on quitting is at most $1/\bar{r}$, then the expected payoff for player i is at most 1, while his payoff is at least 1 if he follows $\sigma^i[h]$. (Recall our assumption that all quitting payoffs are at least 1.) Notice that whenever $(\sigma^i, \pi^{\varepsilon,-i})$ deviates from π^ε, it is by a nonquitting action. Moreover, $(\sigma^i, \pi^{\varepsilon,-i})$ leads to quitting eventually with probability 1.

Because

$$\phi^i(\sigma^i, \pi^{\varepsilon,-i}) \ge \phi^i(\tilde{\sigma}^i, \pi^{\varepsilon,-i}),$$

it suffices to show

$$\phi^i(\sigma^i, \pi^{\varepsilon,-i}) \le \phi^i(\pi^{\varepsilon,i}, \pi^{\varepsilon,-i}) + \varepsilon \tag{2}$$

to prove (1). For any $m \in \mathbb{N} \cup \{0\}$, let σ_m^i be the modification of σ^i, which does not deviate from $\pi^{\varepsilon,i}$ anymore if game plan v^{m+1} becomes active. This means that σ_m^i deviates at most m times. Note that $\sigma_0^i = \pi^{\varepsilon,i}$. Let d_m denote the expected payoff for player i with respect to $(\sigma_m^i, \pi^{\varepsilon,-i})$ and initial state s, i.e.,

$$d_m = \phi^i(\sigma_m^i, \pi^{\varepsilon,-i}).$$

Because $(\sigma^i, \pi^{\varepsilon,-i})$ from initial state s leads to quitting eventually, with probability 1, we must have

$$\phi^i(\sigma^i, \pi^{\varepsilon,-i}) = \lim_{m\to\infty} \phi^i(\sigma_m^i, \pi^{\varepsilon,-i}) = \lim_{m\to\infty} d_m.$$

Let $H(1) \subseteq H^{v^1}$ denote the set of histories h such that (1) player i controls the final state of h; say, u, (2) $\sigma^i(h)$ prescribes to deviate by playing some action $a \in A_u$. For $h \in H(1)$, let $\tau(1,h)$ denote the event that h occurs. The construction of σ^i guarantees that a is nonquitting, hence some game plan v^2 will be chosen after the deviation. With $\mathbb{E}$ denoting the expectation with respect to $(\sigma^i, \pi^{\varepsilon,-i})$, we have by the choice of v^2

$$\begin{aligned}
\mathbb{E}(\phi^i(v^2)|\tau(1,h)) &\le \sum_{s^2 \in S_u(a)} p_u(a,s^2)\beta_u(s^2,\alpha^{k+1}) + \varepsilon/4 \\
&= \gamma_u(a,\alpha^{k+1}) + \varepsilon/4 \\
&\le \delta_u(\alpha^{k+1}) + \varepsilon/4 \\
&= \alpha_u^{k+2} + \varepsilon/4 \\
&\le \alpha_u^k + (\alpha_u^{k+2} - \alpha_u^k) + \varepsilon/4 \\
&\le \phi^i(v^1[h]) + \|\alpha^{k+2} - \alpha^k\| + \varepsilon/4,
\end{aligned}$$

where the last inequality follows from the viability of v^1 with respect to α^k. Thus, with $\mathbb{P}$ denoting the probability of an event with respect to $(\sigma^i, \pi^{\varepsilon,-i})$, we obtain

$$d_1 - d_0 = \sum_{h \in H(1)} \{\mathbb{P}(\tau(1,h)) \cdot [\mathbb{E}(\phi^i(v^2) \mid \tau(1,h)) - \phi^i(v^1[h])]\} \le \|\alpha^{k+2} - \alpha^k\| + \frac{\varepsilon}{4}.$$

In a similar fashion,

$$d_2 - d_1 \le \|\alpha^{k+3} - \alpha^{k+1}\| + \frac{\varepsilon}{8},$$

and, in general,

$$d_{m+1} - d_m \le \|\alpha^{k+m+2} - \alpha^{k+m}\| + \frac{\varepsilon}{2^{m+2}}.$$

Hence

$$d_m - d_0 \le \sum_{l=k}^{k+m-1} \|\alpha^{l+2} - \alpha^l\| + \sum_{l=2}^{m+1} \frac{\varepsilon}{2^l} \le \sum_{l=k}^{\infty} \|\alpha^{l+2} - \alpha^l\| + \frac{\varepsilon}{2}.$$

The choice of k implies

$$\begin{aligned}\sum_{l=k}^{\infty} \|\alpha^{l+2} - \alpha^l\| &\le \sum_{l=k}^{\infty}\sum_{u\in S}(\alpha_u^{l+2} - \alpha_u^l)\\ &= \sum_{u\in S}\sum_{l=k}^{\infty}(\alpha_u^{l+2} - \alpha_u^l)\\ &= \sum_{u\in S}[(\alpha_u^* - \alpha_u^k) + (\alpha_u^* - \alpha_u^{k+1})] \le \frac{\varepsilon}{2}.\end{aligned}$$

Then

$$\phi_s^i(\sigma^i, \pi^{\varepsilon,-i}) = \lim_{m\to\infty} d_m \le d_0 + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \phi_s^i(\pi^\varepsilon) + \varepsilon,$$

which completes the proof of (2).

Deterministic Transitions: Now, suppose, additionally, that all transitions in the game G are deterministic. Then, every quitting game plan induces one specific history, which ends when a specific player quits. Thus, quitting game plans can only induce finitely many different payoffs. As a consequence, in the definition of $\beta_s(t, \alpha^*)$, the infimum is attained for all $s, t \in S$. Moreover, $\delta_t(\alpha^*) = \alpha_t^*$ for all $t \in S$, since $\alpha_t^* = \alpha_t^{k+1} = \delta(\alpha^k) = \delta(\alpha^*)$ for k sufficiently large. For these reasons, the proof for the general case can be applied with $\varepsilon = 0$. □

Example 4. Consider the following example with three players and three states. Player i controls state i and has three actions; namely, one quitting action and two nonquitting actions that lead to states $i+1$ and $i+2$, respectively (where 4 and 5 correspond to states 1 and 2, respectively). If player 1 quits, the payoff is $(2, 1, 4)$; if player 2 quits, the payoff is $(4, 2, 1)$, and if player 3 quits, the payoff is $(1, 4, 2)$. It is shown in Kuipers et al. [7] that this game does not admit a stationary 0-equilibrium, and one can similarly verify that it does not admit stationary ε-equilibria either for small $\varepsilon > 0$.

Our iterative method yields $\alpha^k = \alpha^* = (2, 2, 2)$ for all k. Recall from Example 3 that a quitting game plan for initial state $t_1 \in S$ is denoted by $(t_1, t_2, \ldots, t_m, *)$. We obtain the following viable game plans for state 1 with respect to α^*: $V_1(\alpha^*) = \{(1, *), (1, 2, *)\}$.[5] For the other two states, we have $V_2(\alpha^*) = \{(2, *), (2, 3, *)\}$ and $V_3(\alpha^*) = \{(3, *), (3, 1, *)\}$. Therefore, a subgame-perfect 0-equilibrium for initial state 1 is as follows. Player 1 is supposed to execute game plan $(1, *)$, i.e., to quit. If player 1 deviates by playing the action, which leads to state 2, then the players are supposed to execute game plan $(2, 3, *)$, which minimizes player 1's payoff among the game plans in $V_2(\alpha^*)$. On the other hand, if player 1 deviates by playing the action, which leads to state 3, then player 3 is supposed to execute game plan $(3, *)$. Any further deviations are countered in a similar fashion.

Concluding remarks. **(1) Pure subgame-perfect 0-equilibria.** In all examples we have analyzed, we found a pure subgame-perfect 0-equilibrium. Whether or not this holds, in general, is unclear. Nevertheless, in every game in $\mathcal{G}^+$ for which (1) δ has a fixed point $\bar{\alpha}$, i.e., $\delta(\bar{\alpha}) = \bar{\alpha}$ and (2) the infimum is attained in the definition of the function β with respect to $\bar{\alpha}$, the existence of a pure subgame-perfect 0-equilibrium follows. We do not know if all games in $\mathcal{G}^+$ satisfy these properties, but as we have shown, all games in $\mathcal{G}^+$ with deterministic transitions do have these properties, and consequently, they possess subgame-perfect 0-equilibria. (See also the discussion in §2).

(2) Whether or not α^* is a fixed point of δ. One can prove that $V_t(\alpha^*) \neq \varnothing$ for all $t \in S$. The reason is that the constructed game plans in the proof of Lemma 8 are all stationary. Thus, for a given state t, one can choose stationary game plans $v_t^k \in V_t(\alpha^k)$ for all $k \ge 0$. Since there are only finitely many stationary game plans, it follows that $(v_t^k)_{k\in\mathbb{N}}$ contains a constant subsequence, and it is straightforward to prove that this constant is a stationary game plan in $V_t(\alpha^*)$. Even though $V_t(\alpha^*) \neq \varnothing$ holds for all $t \in S$, it remains unclear whether or not $\delta(\alpha^*) = \alpha^*$. So, it also remains open if δ has a fixed point at all.

[5] For simplicity, we leave out all game plans of the form $(1, 2, 1, 2, \ldots, 1, 2, *)$ other than $(1, 2, *)$.

(3) A polynomial time algorithm for deterministic transitions. In the case of deterministic transitions, a polynomial time algorithm exists to determine the vector α^*. Also, the game plans in a subgame-perfect equilibrium can be determined during play when needed, in polynomial time. To see this, note that for every $t \in S$, the number α_t^k can only have $|S|$ different values. Since every such number is nonincreasing, it follows that the vector α^k can change at most $(|S|-1)|S|$ times before $\alpha^{k+1} = \alpha^k$. Hence the calculation of α^* requires at most $(|S|-1)|S|$ iterations. As one iteration requires the calculation of $|S|^2$ numbers $\beta_s(t, \alpha^k)$, it suffices to show that $\beta_s(t, \alpha)$ can be calculated in polynomial time. Since transitions are deterministic, we have

$$\beta_s(t, \alpha) = \min\{r_u^{i_s} \mid v \in V_t(\alpha) \text{ exists that terminates in } u\}.$$

Hence the calculation of $\beta_s(t, \alpha)$ can be done by a check for every $u \in S$ whether a viable game plan for t with respect to α exists that terminates in u. To do the check, construct the digraph with vertex set

$$V = \{u' \in S \mid \alpha_{u'} \le r_u^{i_{u'}}\}$$

and arc set

$$A = \{(u', u'') \mid \text{an action } a \in A_{u'} \text{ exists such that } p_{u'}(a, u'') = 1\}.$$

Now, observe that a viable game plan for t with respect to α terminating in u exists if and only if the digraph (V, A) has a directed path from t to u. Because the construction of each digraph and the detection of a directed path can be done in polynomial time, it follows that the calculation of $\beta_s(t, \alpha)$ requires polynomial time.

(4) Uniformity. Consider a game $G \in \mathcal{G}^+$ with some initial state s. Let G^m denote the game, which is identical to G, except that each player's reward is the average of his payoffs during the first m stages.[6] (One could also think of G^m as a game in which play ends after stage m.) We claim that, for any $\varepsilon > 0$, the joint strategy π^ε constructed in the proof of the Main Theorem is not only an ε-equilibrium in the infinite game G, but is also a 2ε-equilibrium in the game G^m, given m is sufficiently large. We argue as follows. For any $m \in \mathbb{N}$, let $\phi_s^{m,i}$ denote the reward for player i in game G^m. Since π^ε leads to absorption in G with probability 1, we have that $\phi_s^{m,i}(\pi^\varepsilon)$ converges to $\phi_s^i(\pi^\varepsilon)$ as m tends to infinity. Hence there exists an M such that $|\phi_s^{m,i}(\pi^\varepsilon) - \phi_s^i(\pi^\varepsilon)| \le \varepsilon$ holds for all $m \ge M$. We now prove that π^ε is a 2ε-equilibrium in G^m for all $m \ge M$. Take some $m \ge M$. Consider for player i an arbitrary pure deviation τ^i from the strategy $\pi^{\varepsilon,i}$. Notice that $\phi_s^{m,i}(\tau^i, \pi^{\varepsilon,-i}) \le \phi_s^i(\tau^i, \pi^{\varepsilon,-i})$ as all absorbing payoffs are nonnegative (with regard to any play, payoffs from stage $m+1$ onward are never lower than the payoffs up to stage m). Hence

$$\phi_s^{m,i}(\tau^i, \pi^{\varepsilon,-i}) \le \phi_s^i(\tau^i, \pi^{\varepsilon,-i}) \le \phi_s^i(\pi^\varepsilon) + \varepsilon \le \phi_s^{m,i}(\pi^\varepsilon) + 2\varepsilon,$$

where the second inequality follows from the fact that π^ε is an ε-equilibrium in G. This means that π^ε is a 2ε-equilibrium in G^m indeed.

Of course, after any history h, we can choose an M^h such that the continuation strategy $\pi^\varepsilon[h]$ is a 2ε-equilibrium in $G^m[h]$ given $m \ge M^h$. Our construction does not guarantee, however, that we can choose an M independent of h. For the special case of deterministic transitions, however, such an M exists. The reason is that, as already discussed for games with deterministic transitions, each game plan leads to absorption in one state with probability 1, and therefore it suffices to consider finitely many different game plans (cycles before absorption can be left out).

(5) Negative payoffs. Our method strongly relies on the fact that, if all the payoffs are nonnegative, there exists a subgame-perfect ε-equilibrium for every $\varepsilon > 0$, which is both pure and absorbing. However, when there are negative payoffs, the players may not have an incentive to quit, and pure and absorbing subgame-perfect ε-equilibria may fail to exist. Consequently, our method is no longer applicable in its present form to negative payoffs. For a game without pure subgame-perfect ε-equilibria, we refer to Solan and Vieille [20]. The following game, with two players and two nonabsorbing states, demonstrates that absorbing subgame-perfect ε-equilibria may also fail to exist: Player 1 controls state 1 and has two actions. His first action leads to state 2 with probability 1, whereas his second action, a quitting action, yields a payoff of $(-2, -1)$. Player 2 controls state 2. His first action brings the game to state 1 with probability 1 and his second action is a quitting action, resulting in a payoff of $(-1, -2)$. It is obvious that, for any $\varepsilon \in [0, 1)$, the unique ε-equilibrium is to never quit, which yields a payoff of $(0, 0)$. Note that our iterative method yields $\alpha^k = \alpha^* = (-2, -2)$ for all k.

[6] We are considering the original interpretation of games in $\mathcal{G}^+$, as in §2, when play is infinite (and no quitting is possible). We, however, do assume that the controlling players in the absorbing states have only one action (see point (2) of the reduction in §3).

(6) Games without perfect information. When we allow more than one player to choose actions simultaneously in every state, the situation changes drastically. Take, for instance, the following two-player zero-sum game with three states. In state 1, both players have two actions: a and b for player 1, and c and d for player 2. If the pair of actions chosen by the players is (a, c) or (b, d), then a transition occurs to state 2 with probability 1, whereas if the pair of actions is (a, d) or (b, c), then a transition occurs to state 3 with probability 1. State 2 is absorbing with payoffs $(1, -1)$, whereas state 3 is absorbing with payoff $(-1, 1)$. Notice that the game with initial state 1 is strategically equivalent with a one-shot game, and the unique equilibrium is a mixed one in which both players choose both actions with probability $\frac{1}{2}$. Because we have only examined pure game plans, it is not completely clear if our iterative method can be extended to games without perfect information.

Acknowledgments. The authors thank Eilon Solan, the associate editor, and the referees for their helpful suggestions, which helped us improve the presentation of this paper.

References

[1] Aumann, R. J. 1964. Mixed and behavior strategies in infinite extensive games. M. Dresher, L. S. Shapley, A. W. Tucker, eds. *Advances in Game Theory*. Princeton University Press, Princeton, NJ, 627–650.
[2] Blackwell, D., T. S. Ferguson. 1968. The big match. *Ann. Math. Statist.* **39** 159–163.
[3] Flesch, J., G. Schoenmakers, O. J. Vrieze. 2008. Stochastic games on a product state space. *Math. Oper. Res.* **33**(1) 403–420.
[4] Flesch, J., G. Schoenmakers, O. J. Vrieze. 2009. Stochastic games on a product state space: The periodic case. *Internat. J. Game Theory* **38** 263–289.
[5] Flesch, J., F. Thuijsman, O. J. Vrieze. 2007. Stochastic games with additive transitions. *Eur. J. Oper. Res.* **179** 483–497.
[6] Gillette, D. 1957. Stochastic games with zero stop probabilities. M. Dresher, A. W. Tucker, P. Wolfe, eds. *Contributions to the Theory of Games III*. Princeton University Press, Princeton, NJ, 179–187.
[7] Kuipers, J., J. Flesch, G. Schoenmakers, O. J. Vrieze. 2009. Subgame-perfect equilibria in free transition games. *Eur. J. Oper. Res.* **199** 442–447.
[8] Maitra, A., W. Sudderth. 1993a. Finitely additive and measurable stochastic games. *Internat. J. Game Theory* **22** 201–223.
[9] Maitra, A., W. Sudderth. 1993b. Borel stochastic games with limsup payoff. *Ann. Probab.* **21** 861–885.
[10] Mashiah-Yaakovi, A. 2008. Subgame perfect equilibria in stopping games. Working paper, Tel Aviv University, Tel Aviv, Israel.
[11] Mashiah-Yaakovi, A. 2009. Periodic stopping games. *Internat. J. Game Theory* **38** 169–181.
[12] Mertens, J. F., A. Neyman. 1981. Stochastic games. *Internat. J. Game Theory* **10** 53–66.
[13] Monash, C. A. 1980. Stochastic games: The minimax theorem. Doctoral dissertation, Harvard University, Cambridge, MA.
[14] Shmaya, E., E. Solan. 2004. Two player non zerosum stopping games in discrete time. *Ann. Probab.* **32** 2733–2764.
[15] Shmaya, E., E. Solan, N. Vieille. 2003. An application of Ramsey theorem to stopping games. *Games Econom. Behavior* **42** 300–306.
[16] Simon, R. S. 2003. The structure of non-zero-sum stochastic games. *Adv. Appl. Math.* **38** 1–26.
[17] Solan, E. 1999. Three-player absorbing games. *Math. Oper. Res.* **24** 669–698.
[18] Solan, E. 2005. Subgame-perfection in quitting games with perfect information. *Math. Oper. Res.* **30**(1) 51–72.
[19] Solan, E., N. Vieille. 2001. Quitting games. *Math. Oper. Res.* **26**(2) 265–285.
[20] Solan, E., N. Vieille. 2003. Deterministic multi-player Dynkin games. *J. Math. Econom.* **39** 911–929.
[21] Sorin, S. 1986. Asymptotic properties of a non-zerosum game. *Internat. J. Game Theory* **15** 101–107.
[22] Thuijsman, F., T. E. S. Raghavan. 1997. Perfect information stochastic games and related classes. *Internat. J. Game Theory* **26** 403–408.
[23] Vartiainen, H. 2008. One-deviation principle in coalition formation. Working paper, Aboa Centre for Economics 35, Finland.
[24] Vieille, N. 2000a. Equilibrium in 2-person stochastic games I: A reduction. *Israel J. Math.* **119** 55–91.
[25] Vieille, N. 2000b. Equilibrium in 2-person stochastic games II: The case of recursive games. *Israel J. Math.* **119** 93–126.